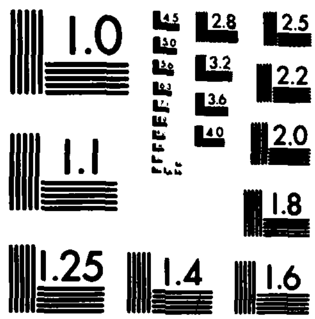
AD-A124 368 CONVERGENCE RATES FOR MULTIVARIATE SMOOTHING SPLINE FUNCTIONS(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER D D COX OCT 82 MRC-TSR-2437 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END
DATE
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2437

CONVERGENCE RATES FOR MULTIVARIATE SMOOTHING SPLINE FUNCTIONS

Dennis D. Cox

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

October 1982

(Received March 5, 1982)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

CONVERGENCE RATES FOR MULTIVARIATE *SMOOTHING SPLINE FUNCTIONS*

Dennis D. Cox*

Technical Summary Report #2437
October 1982

ABSTRACT

Given data $z_i = g(t_i) + \varepsilon_i$, $1 \leq i \leq n$, where g is the unknown function, the t_i are unknown d-dimensional variables in a domain Ω, and the ε_i are i.i.d. random errors, the smoothing spline estimate $g_{n\lambda}$ is defined to be the minimizer over h of $n^{-1}\Sigma(z_i-h(t_i))^2 + \lambda J_m(h)$, where $\lambda > 0$ is a smoothing parameter and $J_m(h)$ is the sum of the integrals over Ω of the squares of all the m^{th} order derivatives of h. Under the assumptions that Ω is bounded and has a smooth boundary, $\lambda \to 0$ appropriately, and the t_i become dense in Ω as $n \to \infty$, bounds on the rate of convergerence of the expected square of $p^{\underline{th}}$ order Sobolev norm (L_2 norm of p^{th} derivatives), are obtained. These extend known results in the one dimensional case. The method of proof utilizes an approximation to the smoothing spline based on a Green's function for a linear elliptic boundary value problem. Using eigenvalue approximation techniques, these rate of convergence results are extended to faily arbitrary domains including $\Omega = R^d$, but only for the case $p = 0$, i.e. L_2 norm.

AMS (MOS) Subject Classifications: Primary 62G05, Secondary 62J99, 41A15, 41A25.

Key Words: Smoothing spline, nonparametric regression, rates of convergence, Sobolev norms, elliptic boundary value problems.

Work Unit Number 4 (Statistics and Probability)

---

*Department of Statistics, 1210 W. Dayton St, University of Wisconsin-Madison, Wisconsin 53706.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 1.

SIGNIFICANCE AND EXPLANATION

Smoothing splines are used to approximate smooth functions when there are only available noisy values of the function at discrete values of the independent variables. It is shown herein that as the grid of values of the independent variables becomes denser in the region of interest, the smoothing spline estimate approaches the true function. Results on the rate of this convergence are given. Convergence of derivatives is investigated, also, but under the assumption that the region is bounded. The theory of linear elliptic partial differential equations is used extensively, along with eigenvalue approximation methods.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

CONVERGENCE RATES FOR MULTIVARIATE
SMOOTHING SPLINE FUNCTIONS

Dennis D. Cox*

1. Introduction.

Let Ω be a domain in d-dimensional Euclidean space $\mathbb{R}^d$. Suppose that there is an unknown real valued function g defined on Ω for which the following data is available

$$z_k = g(t_k) + \varepsilon_k \quad k = 1,2,\ldots,n \tag{1.1}$$

where $t_1,t_2,\ldots,t_n$ are known points in Ω (referred to as "knots"), and $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ are (unknown) random errors. We would like to obtain an estimate of g from the data vector $\underset{\sim}{z} \in \mathbb{R}^n$. If it is known *a priori* that g is a smooth function, then a smoothing spline will provide a reasonable estimate of g. These have been widely used for such estimation problems (see e.g. Wahba [21] or Ragozin, et. al. [10]), and can be justified statistically as either optimal Bayes estimates (Kimeldorf and Wahba [7]) or pointwise minimax estimates (Speckman [13]). In order to describe these estimates, some notation is needed.

A multi-index $\alpha \in \mathbb{Z}_+^d$ is a d-vector whose coordinates belong to $\mathbb{Z}_+ = \{0,1,2,\ldots\}$. The order of α, denoted $|\alpha|$, is given by

$$|\alpha| = \sum_{j=1}^{d} \alpha_j .$$

Given $\alpha \in \mathbb{Z}_+^d$ with $|\alpha| = m$, there is an associated partial differentiation operator of order m given by

$$D^\alpha = \prod_{j=1}^{d} \frac{\partial^{\alpha_j}}{\partial t_j^{\alpha_j}} .$$

*Department of Statistics 1210 W. Dayton St, University of Wisconsin-Madison, Wisconsin 53706.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 1.

For any domain $\Omega \subseteq \mathbb{R}^d$, let $W_2^m(\Omega)$ for $m = 0,1,\ldots$ denote the Sobolev space of all generalized functions (Schwarz's distributions) for which all partial derivatives of order $\leq m$ exist in $L_2(\Omega)$. (For a more formal definition, consult Adams [1], Agmon [2], or Triebel [17].) Letting $\langle\cdot,\cdot\rangle_0$ denote the usual inner product on $L_2(\Omega) = W_2^0(\Omega)$, we define the inner product of $W_2^m(\Omega)$ by

$$\langle f,g\rangle_m = \sum_{|\alpha|\leq m} \langle D^\alpha f, D^\alpha g\rangle_0 ,$$

where the sum is over all multi-indices α of order $\leq m$. The associated norm is

$$\|f\|_m = (\langle f,f\rangle_m)^{1/2} .$$

With these definitions, $W_2^m(\Omega)$ is a Hilbert space.

A related semi-inner product is

$$(f,g)_m = \sum_{|\alpha|=m} \langle D^\alpha f, D^\alpha g\rangle_0 ,$$

which gives rise to the seminorm

$$|f|_m = ((f,f)_m)^{1/2} .$$

A norm equivalent to $\|\cdot\|_m$ which is often convenient is

$$|||f|||_m = (|f|_0^2 + \sum_{k=1}^{d} |D_k^m f|_0^2)^{1/2} , \tag{1.2}$$

where D_k is the operator for partial differentiation in the k^{th} variable. See Theorem 4.2.4 of Triebel [17]. We shall frequently need other norms, in which case the space on which the norm is defined and finite will appear as a subscript to the norm symbol.

The smoothing spline estimate of g can now be described. First make the following

<u>Assumption 1.</u> The data vector $\underset{\sim}{z}$ is given by (1.1) where $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ are uncorrelated random variables with mean 0 and positive, finite variance σ^2.

<u>Definition 1.1.</u> Let $m > d/2$ be the order, $\Delta_n = \{t_1,t_2,\ldots,t_n\} \subseteq \Omega$ be a set of knots, and $\lambda > 0$ be a smoothing parameter. Then the smoothing spline operator $S_{n\lambda}: \mathbb{R}^n \to W_2^m(\Omega)$ (if it exists) is given for $\underset{\sim}{z} \in \mathbb{R}^n$ by

$$S_{n\lambda}\underline{z} = g_{n\lambda} \text{ if and only if compared}$$

$$\text{to any } h \in W_2^m(\Omega)\ , \ g_{n\lambda} \text{ minimizes}$$

$$L_{n\lambda}(h) = n^{-1}\sum_{k=1}^{n}(z_k - h(t_k))^2 + \lambda|h|_m^2 \ .$$

<u>Remarks 1.2</u>(i) The existence of a unique such $g_{n\lambda}$ is assured if the knot set Δ_n is m-unisolvent, as shown in Proposition 2.1 below. A set of points $\Delta_n \subseteq \mathbb{R}^d$ is called m-unisolvent if for any polynomial ψ on $\mathbb{R}^d$ of degree $\leq m - 1$, the condition $\psi(t_k) = 0$ for all k, $1 \leq k \leq n$, implies $\psi(t) = 0$ for all t.

(ii) The choice of the smoothing parameter λ is a problem of considerable interest which we do not discuss here. See Craven and Wahba [5] for one popular method.

(iii) The choice of m is generally dictated by one's prior knowledge of the smoothness properties of g, or by the use that one will make of the estimate $g_{n\lambda}$, as in Ragozin, et. al. [10]. We require $m > d/2$ in order that the evaluations $h(t_i)$ be well defined.

(iv) For a discussion and references on computational aspects, consult Wendelberger [24]. #

Now our main interest here is in showing that $g_{n\lambda}$ *converges to* g in various norms as $n \to \infty$, provided that g satisfies certain conditions, and that λ and Δ_n vary with n in an appropriate way. For $d = 1$, there have been results presented by numerous authors (Wahba [21], Craven and Wahba [5], Utreras [18] [19], Speckman [14], Ragozin [9], and Cox [4]), but for $d > 1$ there have so far been only conjectures (Wahba [22]). The results presented herein essentially generalize all the previous results for $d = 1$ (Theorem 5.1). The main result states that as $n \to \infty$,

$$E\|g_{n\lambda} - g\|_p^2 = \lambda^{-p/m}O[\lambda^{q/m} + n^{-1}\lambda^{-d/2m}] \ .$$

The requirements, briefly, are that $\lambda \to 0$ appropriately as $n \to \infty$. (Assumption 2 below), that Δ_n becomes uniformly dense in Ω (Assumption 3), and that Ω be bounded with smooth boundary (Assumption 4). The number q appearing in the bound is determined

from g according to its differentiability and satisfaction of certain boundary conditions. These are spelled out in Remarks 5.6. If $g \in W_2^m(\Omega)$, then we may take $q = m$.

The smoothing splines which are used most frequently are computed for $\Omega = \mathbb{R}^d$, to which the result described above does not apply. However, we show in Theorem 6.2 that if Δ_n becomes dense in a subdomain of $\mathbb{R}^d$ satisfying the above assumptions, then

$$E\left[n^{-1}\sum_{k=1}^{n}(g_{n\lambda}(t_k) - g(t_k))^2\right] = \mathcal{O}(\lambda + n^{-1}\lambda^{-d/2m}) ,$$

provided $g \in W_2^m(\Omega)$. If we put $\lambda = n^{-2m/(2m+d)}$, then the upper bound becomes $\mathcal{O}(n^{-2m/(2m+d)})$, which is the conjectured optimal rate of convergence in [22]. We note that for $d = 1$, it is immaterial whether $\Omega = \mathbb{R}$ or Ω is some finite open interval containing the knots, since the two minimization problems lead to the same result. This is why the stronger results hold for the unidimensional case.

In order to present our assumptions, further notation and terminology is required. Let $F_n(t)$, $t \in \mathbb{R}^d$ denote the cumulative distribution function (c.d.f.), of the probability measure which assigns mass n^{-1} to the point t_k for each $t_k \in \Delta_n$, i.e.

$$F_n(t) = \sum_{k: t_k \le t} n^{-1}$$

Here, the sum in the numerator is over all k, $1 \le k \le n$, for which all coordinates of t_k are $\le$ the corresponding coordinates of t. Our basic assumption concerning the knots is that the sequence $\{F_n\}$ converges uniformly (at a certain rate) to a c.d.f. F satisfying certain conditions. Our notation suggests that $\Delta_n \subseteq \Delta_m$ if $n \le m$, but in fact it is only the quantity

$$d_n = \sup_t | F(t) - F_n(t) |$$

which is important asymptotically. Our main requirements for λ, Δ_n, and Ω are the following:

<u>Assumption 2.</u> Suppose there is a sequence of intervals $\{[\lambda_n, \Lambda_n]\}$ ($\lambda_n \le \Lambda_n$ for all n) such that $\lambda \in [\lambda_n, \Lambda_n]$ for all n sufficiently large, and that

$$\lim_{n\to\infty} d_n \lambda_n^{-5d/4m} = \lim_{n\to\infty} \Lambda_n = 0$$

Assumption 3. Suppose the limiting knot distribution F has a density $f \in C^\infty(\bar{\Omega})$ with respect to d-dimensional Lebesgue measure such that for all $t \in \Omega$

$$0 < \kappa_1 \le f(t) \le \kappa_2 < \infty$$

for some constants κ_1, κ_2.

Assumption 4. Suppose Ω is a bounded simply connected open domain and its boundary $\partial\Omega$ is C^∞, i.e. there exists a finite open covering $\{O_i\}_{i=1}^{j}$ of $\partial\Omega$ and infinitely differentiable bijections $f_i : O_i \to \mathbb{R}^d$ for which $f_i(O_i \cap \partial\Omega) \subseteq \{x \in \mathbb{R}^d : x_d = 0\}$.

The following result is an immediate consequence of Theorem 5.1.

Corollary 1.3. Suppose that Assumptions 1 through 4 hold, that $m > 3d/2$, and that $g \in W_2^m(\Omega)$. Then for any $p \in \mathbb{Z}+$ satisfying $p < m$,

$$E\|g_{n\lambda} - g\|_p^2 = O(\lambda^{(m-p)/m} + n^{-1}\lambda^{-(2p+d)/2m}) .$$

Furthermore, if $\lambda = Cn^{-2m/(2m+d)}$ for some constant C, then

$$E\|g_{n\lambda} - g\|_p^2 = O\left(n^{-2(m-p)/(2m+d)}\right)$$

Remarks 1.4(i) We conjecture that the latter rate of convergence is the best possible for general $g \in W_2^m(\Omega)$. Under a somewhat different model, Stone [16] showed that for any fixed $t \in \Omega$, $-[2(m-p)/(2m+d)] \log n$ is the optimal rate of convergence in probability of $\log |\hat{g}_n(t) - g(t)|^2$, where $\hat{g}_n(t)$ is any sequence of estimates of $g(t)$. Besides the fact that our results are global, whereas Stone's are local, Stone uses a model wherein $(t_1,z_1),\ldots,(t_n,z_n)$ are a sequence of independent and identically distributed random vectors, and also makes stronger assumptions on the m^{th} derivative than just $g \in W_2^m(\Omega)$. For $d = 1$ and the same observation model, Speckman [15] has obtained the result that the best possible rate of convergence of

$$\log E\, n^{-1} \sum_{k=1}^{n} [\hat{g}_n(t_k) - g(t_k)]^2$$

is $-[2m/(2m+1)] \log n$, where $\hat{g}_n$ is restricted to the class of linear estimates, and $g \in W_2^m(a,b)$. Arguments given below will show that Speckman's seminorm is asymptotically

equivalent to $\|\cdot\|_0$. Speckman even constructs an estimate which achieves the exact optimal rate of convergence (constants included). His estimate is also a spline function, but one that is in general smoother than the smoothing splines we deal with here.

(ii) If one strengthens or weakens the assumption $g \in W_2^q(\Omega)$ then faster or slower rates of convergence are obtained, but some new complications are introduced. The proper function spaces to use are $N_m^q(\Omega)$ defined in Section 3, which are Besov spaces with boundary conditions (Section 4.3.3 of [17]). For nonnegative integers $q \le m$, we have $N_m^q = W_2^m$, but the spaces N_m^q are defined for all real q, so one must deal with "fractional" order Sobolev spaces. If $q < m + 1/2$ then no boundary conditions are active. However, for $m + 1/2 \le q < 3m + 1/2$, certain "natural" boundary conditions of the form $B_k g = 0$ are required for g to be in $N_m^q(\Omega)$. These are derived in Proposition 2.2. For $3m + 1/2 \le q$, still more boundary conditions come into play. Now if $p < 2m - 3d/2$, if $g \in N_m^q(\Omega)$ where q satisfies $d \le q \le 2m + p$, and $q > p$, and if $\lambda = Cn^{-2m/(2q+d)}$, then

$$E\|g_{n\lambda} - g\|_p^2 = O(n^{-2(q-p)/(2q+d)}) ,$$

and this is the best obtainable rate of convergence according to our results. The difficulties with boundary conditions are spelled out in greater detail in Remarks 5.6 below. From a practical point of view, it seems unlikely that our unknown function g will satisfy any of the boundary conditions. Hence, if the function is very smooth (say $g \in W_2^{m+1}(\Omega)$), then $g \in N_m^q(\Omega)$ for all $q < m + 1/2$, and so we can only slightly improve the rate of convergence given in Corollary 1.3. This difficulty with boundary conditions was first studied by Rice and Rosenblatt [11], for $d = 1$.

While the results presented here are significant in and of themselves, they also have applications to other areas. One can easily see that the results on the generalized cross validation method for choosing λ as presented in Craven and Wahba [5] can be both rigorized and generalized. As discussed below, along the say to proving the main theorem, we develop an approximation to $S_{n\lambda}$ which may prove to be useful numerically when smoothing large data sets ($n > 200$, say). Silverman [12] and others have considered penalized likelihood estimates of probability density functions, and we believe the methods

of this work can be applied to this problem as well. Further statistical ramifications of these results will be treated elsewhere.

We now briefly describe the main ideas used in the proof of Theorem 5.1. First, we define a continuous analog of the smoothing spline.

Definition 1.5. Let $m > d/2$ be the order, $\lambda > 0$ a smoothing parameter, and suppose F is a c.d.f. satisfying Assumption 3. Then the continuous smoothing spline operator $S_\lambda : L_2(\Omega) \to W_2^m(\Omega)$ if given for $\zeta \in L_2(\Omega)$ by $S_\lambda \zeta = g_\lambda$ if and only if compared to any other $h \in W_2^m(\Omega)$, g_λ minimizes $L_\lambda(h) = \int (\zeta(t) - h(t))^2 dF(t) + \lambda |h|_m^2$.

Remark. The existence of g_λ is shown in Proposition 2.2 below.

The idea that S_λ could be used to approximate $S_{n\lambda}$ has occurred to many authors (Cogburn and Davis [3], Utreras [18], Speckman [14], Ragozin [9]), but it was in [4], where a rather explicit perturbation formula for $S_{n\lambda}$ in terms of S_λ was developed. As is shown in Section 4 below, there is a function $G_\lambda: \Omega \times \Omega \to \mathbb{R}$ such that for all $\zeta \in L_2(\Omega)$

(1.3) $$(S_\lambda \xi)(t) = \int_\Omega G_\lambda(t,\tau)\xi(\tau)d\tau$$

Now since $S_{n\lambda}$ is a linear operator (it is obtained by minimizing a quadratic form), it may be represented in the form

$$(S_{n\lambda}\underset{\sim}{z})(t) = n^{-1} \sum_{k=1}^{n} g_{n\lambda k}(t) z_k$$

where

(1.4) $$g_{n\lambda k} = n S_{n\lambda} \underset{\sim}{e}_k ,$$

$\underset{\sim}{e}_k$ being the k^{th} coordinate vector in $\mathbb{R}^d$. Noting that the representation for $S_{n\lambda}$ involves an integral with respect to F_n, which approaches F, it is reasonable to expect that as $n \to \infty$, $g_{n\lambda k}(t)$ approaches

(1.5) $$g_{\lambda k}(t) \equiv G_\lambda(t,t_k)/f(t) = G_\lambda(t_k,t)/f(t) .$$

We will see that even though λ is varying with n, $g_{\lambda k}$ is in fact a very good approximation to $g_{n\lambda k}$, provided λ varies in such a way that consistency is obtained. A precise statement on the accuracy of the approximation of $g_{n\lambda k}$ by $g_{\lambda k}$ is given in Theorem 4.3 below.

Acknowledgements. The author has benefited from many suggestions made by Grace Wahba, and is indebted to Sue Leurgans for pointing out an error in the original manuscript. The

author would also like to thank David Ragozin who provided him with an advance copy of [9].

2. Variational Problems. In this section, we show the existence and uniqueness of the smoothing spline and continuous smoothing spline as defined above. Even more importantly, characterizations of these functions are given which will be used in the proof of the main theorem. For convenience, define the n-vector

$$\underset{\sim}{g}(\Delta_n) = (g(t_1),g(t_2),\dots,g(t_n))' \tag{2.1}$$

for any continuous $g : \Omega \to R$.

Proposition 2.1. Suppose that Δ_n is m-unisolvent, and that $m > d/2$. Then $g_{n\lambda} \equiv S_{n\lambda}\underset{\sim}{z}$ exists for all $\underset{\sim}{z} \in R^n$, and is the unique solution of the following problem:

Find $g \in W_2^m(\Omega)$ so that for all $h \in W_2^m(\Omega)$,

$$\lambda(g,h)_m + n^{-1}\sum_{k=1}^{n} g(t_k)h(t_k) = n^{-1}\sum_{k=1}^{n} z_k h(t_k).$$

Proof. For $g,h \in W_2^m(\Omega)$, define $L : R \to R_+$ by

$$L(u;g,h) = L_{n\lambda}(g+uh).$$

If $g_{n\lambda}$ exists then for $\forall h$

$$0 = \frac{dL}{du}(0;g_{n\lambda},h) = 2\lambda(g_{n\lambda},h)_m - 2n^{-1}\sum_k h(t_k)(z_k-g_{n\lambda}(t_k)), \tag{2.2}$$

so it is necessary that $g_{n\lambda}$ be a solution of the problem. Furthermore,

$$\frac{d^2L}{du^2}(u;g,h) = 2\lambda(h,h)_m + 2n^{-1}\sum_{k=1}^{n} h(t_k)^2$$

$$= 2\left[\lambda \sum_{|\alpha|=m} (D^\alpha h,D^\alpha h)_0 + n^{-1}\sum_{k=1}^{n} h(t_k)^2\right].$$

This last expression is always nonnegative, and equals zero just in case

(2.3) $$D^{\alpha}h = 0 \quad \text{for} \quad \forall\, \alpha \in Z_+^d \quad \text{such that} \quad |\alpha| = m,$$
$$\text{and} \quad h(t_k) = 0 \quad \text{for} \quad 1 \le k \le n.$$

The first requirement implies h is a polynomial of degree $\le m - 1$, and the second implies $h \equiv 0$ by m-unisolvency of Δ_n. Hence, $L_{n\lambda}$ is a strictly convex function on $W_2^m(\Omega)$, and any (the) solution g of (2.2) is the unique minimizer of $L_{n\lambda}$.

We now show that $L_{n\lambda}$ has a minimizer, which must then be the unique solution of the problem. Consider the space $H = R^n \times \prod_{j=1}^{N} L_2(\Omega)$ (all products are Cartesian products) where $N = \{{m+d-1 \atop d}\}$, equipped with the Hilbertian norm

$$\|(\underset{\sim}{z}, h_1, \ldots, h_N)\|_H^2 = n^{-1} \sum_{k=1}^{n} z_k^2 + \lambda \sum_{j=1}^{N} |h_j|_0^2 \,.$$

Define the linear operator $U : W_2^m(\Omega) \to H$ by

$$U(h) = (\underline{h}(\Delta_n), D^{\alpha_1}h, \ldots, D^{\alpha_N}h),$$

where $\alpha_1, \ldots, \alpha_N$ are the N distinct elements of Z_+^d satisfying $|\alpha| = m$. By the argument directly following (2.3), U is injective. We claim that $H_0 \equiv$ Range U is a closed subspace of H. This will complete the proof, since $S_{n\lambda}\underset{\sim}{z}$ is then the inverse image under U of the orthogonal projection of $(\underset{\sim}{z}, 0, 0, \ldots, 0) \in H$ onto H_0.

Suppose $\{h_j : j \in Z_+\} \subseteq W_2^m(\Omega)$ is such that $\{Uh_j\}$ us a Cauchy sequence in H. Now $|\cdot|_m$ is a seminorm on $W_2^m(\Omega)$, but a norm on (the equivalence classes in) $W_2^m(\Omega)/P_m$, where P_m is the space of polynomials of degree $\le m - 1$. Note that $W_2^m(\Omega)/P_m$ is a Hilbert space under $(\cdot,\cdot)_m$. Thus we see that for some $\{p_j\} \subseteq P_m$ and $h \in W_2^m(\Omega)$, $\|h_j - h - p_j\|_m \to 0$. Since U is continuous, we also have that

$$(\underline{h}_j(\Delta_n) - \underline{h}(\Delta_n)) - \underline{p}_j(\Delta_n) \to 0 \quad \text{in} \quad R^n.$$

But by assumption, $\{\underline{h}_j(\Delta_n) - \underline{h}(\Delta_n)\}$ converges in R^n as $j \to \infty$, so that $\{\underline{p}_j(\Delta_n)\}$ does also. Since the mapping $p \to \underline{p}(\Delta_n)$ is an isomorphism between P_m and a (closed) subspace of R^n, it follows that $p_j \to p \in P_m$ as $j \to \infty$. Hence, $h_j \to h + p$ in $W_2^m(\Omega)$, and $Uh_j \to U(h+p)$ in H_0, showing H_0 is closed. #

<u>Proposition 2.2.</u> Suppose Assumptions 3 and 4 hold, and $m > d/2$.

(i) For $\forall\, z \in L_2(\Omega)$, $g_\lambda = S_\lambda z$ exists and is the unique solution to the problem:

Find $g \in W_2^m(\Omega)$ such that for all $h \in W_2^m(\Omega)$, $\lambda(g,h)_m + \int gh\, dF = \int zh\, dF$.

(ii) g_λ is the unique solution to a boundary value problem of the form

$$(\lambda f^{-1}(-\Delta)^m + 1)g = z$$

$$B_j g = 0 \text{ on } \partial\Omega \text{ for } 1 \leq j \leq m.$$

Here, Δ is the Laplacian, f is the density for F, and the B_j, $1 \leq j \leq m$, are linear differential operators of order $m + j - 1$. Furthermore, the boundary value problem is a regular elliptic problem.

Proof. (i) This follows by an argument similar to the one used in the proof of Proposition 2.1. Indeed, the argument is somewhat simpler since under Assumptions 3, and 4,

$$\lambda|h|_m^2 + \int h^2 dF$$

gives a norm on $W_2^m(\Omega)$ which is equivalent to $\|\cdot\|_m$ (see (2) in Theorem 4.2.4 of [17]). Thus, it is much easier to show that the appropriate subspace of $H = L_2(\Omega) \times \prod_{j=1}^{N} L_2(\Omega)$ is closed.

(ii) This follows from standard methods. We briefly recap the argument in Agmon [2], pages 141-143. First note that the bilinear form $B(u,v) = (u,v)_m$ is uniformly strongly elliptic (Definition 7.1 of [2]), since

$$B(u,v) = \sum_{|\alpha|=m} (D^\alpha u, D^\alpha v)_0$$

and for all $\zeta \in \mathbb{R}^d$,

$$\sum_{|\alpha|=m} \zeta^{2\alpha} = \left(\sum_{j=1}^{d} \zeta_j^2\right)^m = \|\zeta\|_{\mathbb{R}^d}^{2m}.$$

Hence, by Theorem 10.2 of [2], there exist linear differential operators $B_j = B_j(x,D)$, $x \in \partial\Omega$, of order $m + j - 1$ $(1 \leq j \leq m)$ such that

$$B(u,v) = (u, Av)_0 + \sum_{j=1}^{m} \int_{\partial\Omega} \frac{\partial^{j-1} u}{\partial n^{j-1}} B_{m-j+1} v \, d\sigma ,$$

where $A = (-\Delta)^m$, and the B_j's are nowhere characteristic for $\partial\Omega$. Here, $\partial/\partial n$ denotes differentiation in the direction of the outward normal to $\partial\Omega$, and $d\sigma$ is the element of area on $\partial\Omega$. If g were a solution to the boundary value problem in the statement of

(ii), then for all $h \in W_2^m(\Omega)$ we have

$$\lambda(g,h)_m + \int gh\,d\Gamma = \lambda((-\Delta)^m g,h)_0 + \int g(t)h(t)f(t)dt$$
$$= (f(z-g),h)_0 + (fg,h)_0$$
$$= (fz,h)_0 = \int zh\,d\Gamma.$$

Hence, g would be a solution to the variational problem stated in part (i). See also Lions & Magenes [8], Section 9.5.

We now verify that the boundary value problem is regularly elliptic by checking the conditions of Definition 5.2.1/4 of Triebel [17]. That $\lambda f^{-1}(-\Delta)^m$ is properly elliptic follows from

$$\lambda f(t)^{-1} \sum_{|\alpha|=m} \zeta^{2\alpha} = \lambda f^{-1} \|\zeta\|_{\mathbb{R}^d}^{2\alpha} > 0 \quad \text{for all} \quad \zeta \in \mathbb{R}^d\backslash\{0\}$$

and, if $\zeta, \eta \in \mathbb{R}^d$ are linearly independent then the roots of

$$P(\tau) = \sum_{|\alpha|=m} (\zeta+\tau\eta)^{2\alpha} = 0$$

are the roots of

$$\sum_{j=1}^{d} (\zeta_j+\tau\eta_j)^2 = 0$$

(replicated m times), and these are clearly complex conjugates of each other. Since the boundary operators $\{B_j\}_{j=1}^m$ are nowhere characteristic for $\partial\Omega$, they form a normal system and furthermore they have orders $\leq 2m-1$. Finally, that B_j satisfies the complementing condition with respect to $\lambda f^{-1}(-\Delta)^m + 1$ follows from Remark 5.2.1/4 of [17]. Hence, the problem is regularly elliptic.

To complete the proof, we need to show that the boundary value problem possesses a unique solution. According to Theorem 5.4.3 of [17] and the remark thereafter, it suffices to show that the only solution to the boundary value problem when $z \equiv 0$ is the trivial solution. However, this is immediate since $\lambda f^{-1}(-\Delta)^m$ is a positive operator so -1 cannot be in its spectrum, i.e. $\lambda f^{-1}(-\Delta)^m + 1$ has empty null space. #

3. Function Spaces. In this section, we will introduce the function spaces which will be used in the sequel. The Sobolev spaces $W_2^k(\Omega)$ were defined in Section 1. It will be useful to have the Besov spaces $B_{22}^s(\Omega)$, which are defined for all real $s > 0$. A complete account may be found in Triebel [17]. The following, brief definition will serve our purposes. Given two Banach spaces A_0, A_1, both being subsets of a larger Banach space A, define the K-functional $K:(0,\infty) \times (A_0 + A_1) \to \mathbb{R}_+$ (here, $A_0 + A_1$ is the linear span of $A_0 \cup A_1$, and x denotes cartesian product) by

$$K(\lambda,a) = \inf \left\{\|a_0\|_{A_0} + \lambda\|a_1\|_{A_1} : a = a_0 + a_1 \text{ and } a_i \in A_i, i = 1,2\right\} .$$

For $\theta \in (0,1)$, the K-method interpolate is the Banach space

$$(A_0,A_1)_{\theta,2} = \{a \in A_0 + A_1 : \|a\|_{(A_0,A_1)_{\theta,2}} \equiv (\int_0^\infty [y^{-\theta}K(\lambda,a)]^2\lambda^{-1}\, d\lambda)^{1/2} < \infty\} .$$

One of the basic properties of this interpolation method is that the spaces are increasing in θ, i.e. if $0 < \theta < \phi < 1$, then $A_0 \cap A_1 \subseteq (A_0,A_1)_{\theta,2} \subseteq (A_0,A_1)_{\phi,2} \subseteq A_1 + A_2$. For any $s \in (0,\infty)$, let k be an integer $> s$. Then we may define the Besov spaces by

$$B_{22}^s(\Omega) = (L_2(\Omega),\ W_2^k(\Omega))_{\theta,2} \quad \text{with } \theta = s/k.$$

The definition is independent of k, up to equivalent renorming. For $s \in Z+$,

$$B_{22}^s(\Omega) = W_2^s(\Omega)$$

with equivalent norms. For $\Omega = R^d$, this result is stated in Remark 2.3.3/4 of [17], and it follows for general Ω from Definition 4.2.1 of that reference. Interpolation of Besov spaces yields nothing new, i.e.

$$(B_{22}^s(\Omega),\ B_{22}^r(\Omega))_{\theta,2} = B_{22}^{(1-\theta)s+\theta r}(\Omega).$$

This follows from Theorem 4.3.1/1 of [17].

Let $N_m(\Omega)$ be the subspace of $W_2^{2m}(\Omega)$ of functions which satisfy the natural boundary conditions $B_j h = 0$ on $\partial\Omega$, $1 \le j \le m$. That these are well defined follows from Sobolev's trace theorem (Theorem 3.10 of Agmon [2]). Also, define the operator

$$\mathcal{D} = f^{-1}(-\Delta)^m$$

with domain of definition $N_m(\Omega)$. Here, Δ is the usual Laplacian on R^d and f^{-1} denotes division by the density of F. According to Theorem 14.6 of Agmon [2], $\mathcal{D}$ has discrete spectrum contained in the positive real axis, and the eigenvalues have finite multiplicity. We will write the eigenvalues as

$$\gamma_1 \leq \gamma_2 \leq \cdots$$

where each is replicated according to its multiplicity. The corresponding eigenfunctions will be denoted $\{\phi_\nu\}_{\nu=1}^{\infty}$, and we may assume that $\{\phi_\nu\}$ forms a complete orthonormal system for $L_2(\Omega)$ (Theorem 16.5 of Agmon [2]). The space N_m is given the following norm

$$\|h\|^2_{N_m(\Omega)} \equiv \sum_\nu (\gamma_\nu^2 + 1)(h,\phi_\nu)_0^2$$

$$= |\mathcal{D}h|_0^2 + |h|_0^2.$$

One notes that $\|\cdot\|_{N_m(\Omega)}$ is the same as defined in 1.18.10 of Triebel [17]. Using the definition contained therein, we may define for any nonnegative real number s,

$$N_m^s(\Omega) = \{h \in L_2(\Omega) : \|h\|_{N_m^s(\Omega)} \equiv \left(\sum_\nu (\gamma_\nu^{s/m} + 1)(h,\phi_\nu)_0^2\right)^{1/2} < \infty\}.$$

According to Theorem 1.18.10 of [17], if $\theta \in (0,1)$,

$$(L_2(\Omega), N_m^s(\Omega))_{\theta,2} = N_m^{\theta s}(\Omega). \tag{3.1}$$

In particular, for $s \in [0,2m]$

$$(L_2(\Omega), N_m(\Omega))_{s/2m,2} = N^s(\Omega).$$

In order to give a more useful characterization of $N_m^s(\Omega)$ it is necessary to introduce Besov spaces with boundary conditions.

Letting B_j be the boundary operators from Proposition 2.2(ii), set

$$B^s_{22,\{B_j\}}(\Omega) = \{h \in B^s_{22}(\Omega) : B_j h|_{\partial\Omega} = 0 \text{ for each } j \text{ satisfying } m + j - 1 < s - 1/2\}.$$

(compare with section 4.3.3 of [17]). Note that

$$N_m(\Omega) = N_m^{2m}(\Omega) = B^{2m}_{22,\{B_j\}}(\Omega).$$

Our immediate aim is to establish similar equivalence results for the spaces $N_m^s(\Omega)$, $s \in R_+$.

Proposition 3.1. Suppose Assumptions 3 and 4 hold.

(i) If $s \in R_+$ and if there exist no i, $1 \leq i \leq m$, and no $j \in Z_+ \backslash \{0\}$ such that $s = 2m(j - 1/2) + i - 1/2$, then

$$N_m^s(\Omega) = B^s_{22,\{C_{ki}\}}(\Omega)$$

Here, the set of boundary operators $\{C_{ki}\}$ are the class of all operators of the form

$$C_{ki} = B_i \circ \mathcal{D}^k \qquad \text{for } 1 \leq i \leq m,\ 1 \leq k \leq j \equiv [s/2m] + 1.$$

(ii) If $s = 2m(j - 1/2) + i - 1/2$ for some i, $1 \leq i \leq m$, and $j \in Z_+ \backslash \{0\}$, then

$$N_m^s(\Omega) = \{h \in B^s_{22,\{C_{kl}\}}(\Omega) : B_i \mathcal{D}^j h \in B_{22}^{1/2}(R^d) \text{ with supp } h \subseteq \bar{\Omega}\}.$$

(iii) For all $s \in R_+$ the norms $\|\cdot\|_{B^s_{22}(\Omega)}$ and $\|\cdot\|_{N^s_m(\Omega)}$ are equivalent on $N_m^s(\Omega)$. In particular, for $k \in Z_+$, the norms $\|\cdot\|_k$ and $\|\cdot\|_{N^k_m(\Omega)}$ are equivalent on $N_m^k(\Omega)$.

__Proof.__ Step 1. First assume $s = 2mk$ for some $k \in Z_+$. We will prove (i) and (iii) in this case by induction on k. The equality in (i) is true for $k = 1$, as already noted. Now we make use of the following __a-priori__ inequality in Theorem 4.3.4. of [17]: for any $p \in Z_+$, there exist positive constants C_1, C_2 such that for all $h \in W_2^{2m+p}(\Omega)$,

$$C_1\|h\|_{2m+p} \leq \|\mathcal{D}h\|_p + \|h\|_0 + \sum_{j=1}^{m} \|B_j h\|_{B_{22}^{m+p+j-1/2}(\Omega)} \leq C_2\|h\|_{2m+p}.$$

Taking $p = 0$, we see that $\|\cdot\|_{N_m^{2m}(\Omega)}$ is equivalent to $\|\cdot\|_{2m}$ on $N_m^{2m}(\Omega)$, since $B_j h = 0$, $\forall\, h \in N_m^{2m}(\Omega)$ and all j, $1 \leq j \leq m$.

For induction step, assume for some $k \geq 2$ that (i) and (iii) hold when $s = 2m(k-1)$. Then $h \in N_m^{2mk}(\Omega)$ implies $\mathcal{D}h \in N_m^{2m(k-1)}(\Omega)$, so that $\mathcal{D}h$ satisfies

$$B_i \mathcal{D}^{(k-1)}(\mathcal{D}h) = B_i \mathcal{D}^k h = 0 \quad \text{on} \quad \partial\Omega$$

for $1 \le i \le m$. Hence, all $h \in N_m^{2mk}(\Omega)$ satisfy the boundary conditions in (i). Now apply the a-priori inequality again to conclude that $\|\cdot\|_{2mk}$ is equivalent to the norm $N_m^{2mk}(\Omega)$ given by

$$\|\mathcal{D}h\|_{2m(k-1)} + \|h\|_0 ,$$

but by induction hypothesis, $\|\mathcal{D}h\|_{2m(k-1)}$ can, for all $h \in N_m^{2mk}(\Omega)$, be trapped between constant multiples of

$$\begin{aligned} \|\mathcal{D}h\|_{N_m^{2m(k-1)}(\Omega)} &= \left(\|\mathcal{D}^{(k-1)}(\mathcal{D}h)\|_0^2 + \|\mathcal{D}h\|_0^2\right)^{1/2} \\ &= \left(\|\mathcal{D}^k h\|_0^2 + \|\mathcal{D}h\|_0^2\right)^{1/2}. \end{aligned}$$

By using the eigenfunction expansion for h, we easily see that

$$\gamma_1^2 \|h\|_0^2 \le \|\mathcal{D}h\|_0^2 \le \gamma_1^{-2(k-1)} \|\mathcal{D}^k h\|_0^2 .$$

It is now clear that $\|\cdot\|_{2mk}$ is equivalent to the norm given by

$$\left(\|\mathcal{D}^k h\|_0^2 + \|h\|_0^2\right)^{1/2} = \|h\|_{N_m^{2mk}(\Omega)} .$$

This shows that (i) and (iii) hold when $s = 2mk$ for some $k \in Z_+$.

Step 2. Parts (i) and (ii) are now an immediate consequence of Step 1 and Theorem 4.3.3. of [17]. Part (iii) follows from the norm equivalences of Step 1 and a simple calculation with the definition of the K-method of interpolation. #

4. <u>The Green's Function Approximation.</u> The main result of this section is Theorem 4.3, which gives an approximation of the smoothing spline in terms of the family of Green's functions descsribed in the introduction. The proof of this theorem depends on the two technical lemmas (Lemmas 4.1 and 4.2). The first provides various norm estimates on the family of Green's functions. The second lemma concerns the accuracy of approximating the inner product in $L_2(F_n)$ by the inner product in $L_2(F)$, and in particular if one of the entries in the inner product is from the Green's function family.

Since $S_\lambda z$ is the solution of a certain boundary value problem (Proposition 2.2(ii)), it has the representation given in (1.3), where $G_\lambda(\cdot,\cdot)$ is the Green's function for $\lambda\mathcal{D}+1$ with domain $N_m(\Omega)$. We prove this as in Theorem 5.6.4 of Triebel [17] by giving an explicit construction for G_λ. Let $\{\gamma_\nu\}$, $\{\phi_\nu\}$ be the eigensystem for the operator $\mathcal{D} = f^{-1}(-\Delta^m)$ with domain $N_m(\Omega)$ as described above. Then we may develop G_λ in a bilinear expansion, viz.

$$G_\lambda(t,\tau) = \sum_\nu (1+\lambda\gamma_\nu)^{-1}\phi_\nu(t)\phi_\nu(\tau),$$

where the series converges (at least) in $L_2(\Omega \times \Omega)$. The aforementioned reference guarantees $G_\lambda \in W_2^p\ (\Omega \times \Omega)$ for any $p \in Z_+$ satisfying $p < 2m-(d/2)$ (and not for any larger p).

We now introduce some notation used in the sequel. For α, $\beta \in Z_+^d$, define

$$G_\lambda^{\alpha,\beta}(t,\tau) = D_t^\alpha D_\tau^\beta G_\lambda(t,\tau)\ ,$$

where D_t denotes differentiation with respect to the first d-dimensional argument, and similarly for D_τ. Also, we shall use the notation for arbitrary nonnegative functions g,h

$$g(t) \approx h(t)$$

to mean there exist positive, finite constants C_1, C_2 such that for all t of interest

$$C_1 g(t) \le h(t) \le C_2 g(t)\ .$$

The notation

$$g(t) \sim h(t) \quad \text{as} \quad t \to t_0$$

means that

$$g(t) - h(t) = o(h(t)) \quad \text{as} \quad t \to t_0 .$$

Finally, we shall need the lin[illegible] [illegible]erator $R_{n\lambda} : C(\Omega) \to W_2^m(\Omega)$ given by

$$(R_{n\lambda}h)(t) = \int G_\lambda(t,\tau)f(\tau)^{-1}h(\tau)dF(\tau) - n^{-1}\sum_{k=1}^{n} G_\lambda(t,t_k)f(t_k)^{-1}h(t_k)$$

$$= \int G_\lambda(t,\tau)f(\tau)^{-1}h(\tau)d[F(\tau) - F_n(\tau)] .$$

<u>Lemma 4.1.</u> Let $2m > d/2$, and let Assumptions 3 and 4 hold.

(i) If $0 \leq p < 2m - d/2$, then

$$\|G_\lambda\|_{W_2^p(\Omega\times\Omega)} \approx \lambda^{-(2p+d)/4m} .$$

Here the constants may be chosen so as to depend only on Ω, m, p, and F.

(ii) If $p \in Z_+$ and $p < 2m-d$, then $G_\lambda \in C_B^p(\Omega \times \Omega)$, the class of functions whose derivatives of order $\leq p$ are bounded and continuous on $\Omega \times \Omega$.

(iii) If $p < 2m-d$, then

$$\sup_t \|G_\lambda(\cdot,t)\|_p \leq K_3\lambda^{-(p+d)/2m}$$

where $K_3 = K_3\ (\Omega,m,p,F)$. Here $\|G_\lambda(\cdot,t)\|_p$ means $\|\cdot\|_p$ computed when G_λ is considered as a function of its first argument only.

<u>Proof.</u> (i) Rather than estimate $\|G_\lambda\|_{W_2^p(\Omega \times \Omega)}$, we work with the equivalent norm

$|||G_\lambda|||_{W_2^p(\Omega\times\Omega)}$ as defined in (1.2). Since $\mathcal{D}$ is formally self adjoint,

$$G_\lambda(t,\tau) = G_\lambda(\tau,t),$$

and we clearly have

$$|||G_\lambda|||^2_{W_2^p(\Omega\times\Omega)} + \|G_\lambda\|^2_{L_2(\Omega\times\Omega)} = 2\int_\Omega |||G_\lambda(\cdot,\tau)|||_p^2 d\tau , \tag{4.1}$$

where by $G_\lambda(\cdot,\tau)$ we mean to consider G_λ as a function of its first argument while the second is held fixed at τ. By Proposition 3.1, there are positive finite constants c_1 and c_2 depending only on Ω,m,p, and f such that for every τ

(4.2) $$c_1 \|G_\lambda(\cdot,\tau)\|_{N_m^p(\Omega)} \le |||G_\lambda(\cdot,\tau)|||_p^2 \le c_2 \|G_\lambda(\cdot,\tau)\|_{N_m^p(\Omega)} .$$

According to Theorem 14.6 of Agmon [2], we have

(4.3) $$\gamma_\nu \approx \nu^{2m/d} .$$

We note that Ω need only satisfy the restricted cone property in order for this to hold (see p. 239 of [2]). Hence, for some constants $C_i = C_i(\Omega,m,p,F)$, $i = 1,2$, it holds that

$$\begin{aligned}\int \|G_\lambda(\cdot,\tau)\|^2_{N_m^p(\Omega)} d\tau &= \sum_\nu (1 + \lambda\gamma_\nu)^{-2}\gamma_\nu^{p/m} \\ &\le C_1 \sum_\nu (1 + C_2\lambda\nu^{2m/d})^{-2}\nu^{2p/d} \\ &\sim C_1\lambda^{-(2p+d)/2m} \int_0^\infty \frac{x^{2p/d}dx}{(1+C_2x^{2m/d})^2}\end{aligned}$$

as $\lambda \downarrow 0$. The last line follows from an application of dominated convergence. This shows that for some finite constant $K_1 = K_1(\Omega,m,p,F)$,

$$\|G_\lambda\|_{W_2^p(\Omega \times \Omega)} \le K_1\lambda^{-(2p+d)/4m} .$$

An entirely similar argument shows

$$\|G_\lambda\|_{W_2^p(\Omega \times \Omega)} \ge K_2\lambda^{-(2p+d)/4m}$$

for some constant $K_2 = K_2(\Omega,m,p,F)$. This completes the proof of (i).

(ii) Since $G_\lambda \in W_2^q(\Omega \times \Omega)$ for any $q < 2m - (d/2)$, it follows that $G_\lambda \in C_B^p(\Omega \times \Omega)$ for any $p < q - (d/2)$ by one of Sobolev's imbedding theorems (see Theorem 5.4, part (C) in Adams [1]).

(iii) Considering $\|G_\lambda(\cdot,t)\|_{N_m^p(\Omega)}$ as a function of t, we have

(4.4) $$\|G_\lambda(\cdot,t)\|^2_{N_m^p(\Omega)} = \sum_\nu (1+\lambda\gamma_\nu)^{-2} \gamma_\nu^{p/m} |\phi_\nu(t)|^2$$

where the series converges (at least) in $L_1(\Omega)$, provided of course that $p < 2m - d/2$. The a priori estimate (Theorem 5.3.4. of Triebel [17]) yields

$$\|\phi_\nu\|_{2m} \le C_3(\|D\phi_\nu\|_0 + \|\phi_\nu\|_0) \le C_3(\gamma_\nu + 1)\ \|\phi_\nu\|_0 = C_3(\gamma_\nu + 1)$$

where the constant $C_3 = C_3(\Omega, m, F)$. By Sobolev's inequality (p. 32 of Agmon [2]), for some $C_4 = C_4(\Omega, m)$ and any $r \ge 1$, we have for all $t \in \Omega$

$$|\phi_\nu(t)| \le C_4(r^{-(2m-d/2)}\ |\phi_\nu|_{2m} + r^{d/2}|\phi_\nu|_0)$$

$$\le C_4\ (C_3\ (\gamma_\nu+1)\ r^{-(2m-d/2)} + r^{d/2}\ ).$$

If we now utilize $\gamma_\nu = \nu^{2m/d}$, and put $r = \nu^{1/d}$, then we obtain $|\phi_\nu(t)| \le C_5\nu^{1/d}$ for some $C_5 = C_5\ (\Omega, m)$. Substituting this back into (4.4) yields

$$\|G_\lambda\ (\cdot,t)\|^2_{W^p_m(\Omega)} \le C_5^2\ \sum_\nu \nu(1+\lambda\gamma_\nu)^{-2}\gamma_\nu^{p/m}$$

$$\le C_5^2\ C_1\ \sum_\nu\ (1+C_2\lambda\nu^{2m/d})^{-2}\nu^{(2p+d)/d}$$

$$\sim C_5^2\ C_1\ \lambda^{-(p+d)/m}\ \int_0^\infty \frac{x^{(2p+d)/d}dx}{(1+C_2x^{2m/d})^2}$$

as $\lambda \downarrow 0$. Note that convergence of the latter integral requires

$$p < 2m - d.$$

In view of Proposition 3.1, the proof of (iii) is complete. ∎

Lemma 4.2.

Let $m > 3d/2$, and let Assumptions 3 and 4 hold.

(i) For functions $h,g \in W_2^d(\Omega)$ we have

$$|\int hg\, d(F-F_n)| \le K_4 d_n \|h\|_d \|g\|_d$$

where $K_4 = K_4(\Omega)$ is a constant, and $d_n = \sup|F-F_n|$.

(ii) Suppose $p \in Z_+$ satisfies $p < 2m - (3d/2)$. Then for any $h \in W_2^d(\Omega)$,

$$\|R_{n\lambda} h\|_p \le K_4' d_n \|G_\lambda\|_{W_2^{p+d}(\Omega\times\Omega)} \|h\|_d ,$$

where $K_4' = K_4\, C$, where C depends only on F, and K_4 is the same constant as in (i).

(iii) Under the same hypotheses as in (ii), for each $\nu \in Z_+$,

$$\|R_{n\lambda}^{\nu} h\|_p \le K_5 (K_6 d_n \lambda^{-5d/4m})^{\nu}\, \lambda^{-(p-d)/2m} \|h\|_d$$

where $K_i = K_i(\Omega,m,p,F)$, $i = 5, 6$, are constants,

(iv) Under the same hypothese as in (ii), there is a constant $K_7 = K_7(\Omega,m,p,F)$ such that

$$n^{-1} \sum_{k=1}^{n} \|g_{\lambda k}\|_p \|g_{\lambda k}\|_d \le K_7 \lambda^{-(p+2d)/2m} (1+d_n \lambda^{-d/m}) ,$$

where $g_{\lambda k}$ is defined in (1.5).

Proof. (i) We start with the following integration by parts formula, valid for any $h \in C_0^\infty(\mathbb{R}^d)$ = infinitely differentiable functions of compact support, and any probability distribution function G on Ω:

(4.5) $$\int h(t)dG(t) = \sum_{\beta\in\{0,1\}^d} (-1)^{|\beta|} \int D^\beta h(t[\beta])\cdot G(t[\beta])dt$$

where $t[\beta]$ has j^{th} coordinate $t_j[\beta]$ given by

$$t_j[\beta] = \begin{cases} t_j & \text{if } \beta_j = 1 , \\ A & \text{if } \beta_j = 0 , \end{cases}$$

where $A > 0$ is chosen so that $\Omega \subseteq [-A,A]^d$ holds. The summation in (4.5) is over all $\beta \in Z_+^d$ whose coordinates are either 0 or 1. The integrations are over $[-A,A]^d$. Note that each integral on the right hand side may be reduced to a $|\beta|$ -dimensional integral with respect to

$$\prod_{j:\beta_j=1} dt_j \ .$$

A proof of (4.5) runs as follows. If G is a unit point mass concentrated at some point in Ω, then (4.5) can be proved by a tedious but straightforward induction on d. The result for arbitrary descrete G follows by taking convex combinations. It then suffices to show that if the formula holds for each element of a sequence $G_1, G_2, \ldots$ which converges to some G uniformly in Ω, then it holds for G, because the discrete probability measures are dense in the space of all probability measures on $[-A, A]^d$ when equipped with the topology induced by Kolmogorov's (sup-norm) metric. Since Kolmogorov's topology is stronger than the topology of weak convergence, we immediately have that

$$\int h(t)dG_n(t) \to \int h(t)dG(t) \ . \tag{4.6}$$

Furthermore, if $\beta \in \{0,1\}^d$ we claim that as $n \to \infty$

$$\int D^\beta h(t[\beta])G_n(t[\beta])dt \to \int D^\beta h(t[\beta])G(t[\beta])dt \ .$$

Since distribution functions are bounded by 1 and $h \in C_0^\infty(R^d)$, this latter result follows by Lebesque's dominated convergence theorem, completing the proof of (4.5).

To complete the proof of (i), assume first of all that $\Omega = [-A,A]^d$ and also that $h,g \in C_0^\infty(R^d)$. Then the product differentiation rule followed by Cauchy-Schwarz yields

$$\begin{aligned} |\int \{(F - F_n)D^\beta hg)\}(t[\beta])dt| &\leq d_n \int |D^\beta(hg)(t[\beta])|dt \\ &\leq d_n \sum_{\alpha\leq\beta} \int | \{(D^\alpha h)(D^{\beta-\alpha} g)\}(t[\beta])|dt \\ &\leq d_n \sum_{\alpha\leq\beta} [\int (D^\alpha h)^2(t[\beta])dt]^{1/2} [\int (D^{\beta-\alpha} g)^2(t[\beta])dt]^{1/2} . \end{aligned} \tag{4.7}$$

If $|\beta| < d$, Sobolev's theorem on traces (p. 38 of Agmon [2]) applies and we have

$$\left(\int (D^{\alpha}h)^2(t[\beta])dt\right)^{1/2} \leq C_1 \|h\|_{W_2^d([-A,A]^d)}$$

where $C_1 = C_1(A)$ is a constant. Hence, from (4.5) and (4.7) we obtain

$$(4.8) \qquad \left|\int hg\, d(F - F_n)\right| \leq 2^{2d} \max\{1,C_1\} d_n \|h\|_{W_2^d([-A,A]^d)} \|g\|_{W_2^d([-A,A]^d)} .$$

This was obtained under the assumptions that $h,g \in C_0^{\infty}(R^d)$ and $\Omega = [-A,A]^d$. However, the restrictions of $C_0^{\infty}(R^d)$ functions to $[-A,A]^d$ gives a dense set in $W_2^d([-A,A]^d)$, so (4.8) holds for arbitrary $h,g \in W_2^d([-A,A]^d)$. Furthermore, there exists a continuous extension operator (Theorem 4.32 of Adams [1]) $E: W_2^d(\Omega) \to W_2^d([-A,A]^d)$, so that $W_2^d([-A,A]^d)$ norms in (4.8) may be replaced by $W_2^d(\Omega)$ norms at the cost of introducing another constant factor which depends on Ω. this completes the proof of (i).

(ii) Define $\bar{G}_{\lambda}(t,\tau) = G_{\lambda}(t,\tau)/f(\tau)$. The assumptions on p guarantee that $\bar{G}_{\lambda}^{\alpha,0}$ is bounded and continuous on $\bar{\Omega} \times \bar{\Omega}$ by Lemma 4.1 (ii) and Assumption 3, provided $|\alpha| \leq p$. Hence, we may interchange differentiations and integrations under these conditions. Thus

$$(4.9) \qquad \|R_{n\lambda}h\|_p^2 = \left\|\int \bar{G}_{\lambda}(\cdot,\tau)h(\tau)\, d[F(\tau) - F_n(\tau)]\right\|_p^2$$

$$= \sum_{|\alpha|\leq p} \left|\int \bar{G}_{\lambda}^{\alpha,0}(\cdot,\tau)h(\tau)\, d[F(\tau) - F_n(\tau)]\right|_0^2$$

$$\leq K_4^2 d_n^2 \|h\|_d^2 \sum_{|\alpha|\leq p} \int \|\bar{G}_{\lambda}^{\alpha,0}(t,\cdot)\|_d^2\, dt ,$$

where the last line follows from part (i). Now

$$\sum_{|\alpha|\leq p} \int \|\bar{G}_{\lambda}^{\alpha,0}(t,\cdot)\|_d^2\, dt = \sum_{|\alpha|\leq p} \sum_{|\beta|\leq d} \int\int |\bar{G}_{\lambda}^{\alpha,\beta}(t,\tau)|^2\, dt d\tau \leq \|\bar{G}_{\lambda}\|_{p+d}^2 .$$

An application of the product differentiation rule, Assumption 3, and Sobolev's inequality yields

$$(4.9a) \qquad \|\bar{G}_{\lambda}\|_{p+d} \leq C\, \|G_{\lambda}\|_{p+d},$$

where C is a constant depending only on f. Substituting this back into (4.9) completes the proof of (ii).

(iii) Using part (ii) and Lemma 4.1 (i) gives for any $\nu \in Z_+ \setminus \{0\}$,

$$\|R_{n\lambda}^{\nu} h\|_p \le K_4' d_n \|G_\lambda\|_{W_2^{p+d}(\Omega\times\Omega)} \|R_{n\lambda}^{\nu-1} h\|_d$$

$$\le C_1 K_4' d_n \lambda^{-(2p+3d)/4m} \|R_{n\lambda}^{\nu-1} h\|_d$$

for some constant $C_1 = C_1(\Omega,m,p)$. Using $p = d$ in this last inequality and iterating gives the desired result.

(iv) We have by Cauchy-Schwarz

$$n^{-1} \sum_{k=1}^{n} \|g_{\lambda k}\|_p \|g_{\lambda k}\|_d = \int \|\bar{G}_\lambda(\cdot,t)\|_p \|\bar{G}_\lambda(\cdot,t)\|_d \, dF_n(t) \tag{4.10}$$

$$\le \left[\int \|\bar{G}_\lambda(\cdot,t)\|_p^2 \, dF_n(t)\right]^{1/2} \left[\int \|\bar{G}_\lambda(\cdot,t)\|_d^2 \, dF_n(t)\right]^{1/2} .$$

Now concentrating on the first factor, we obtain from part (i) that

$$\int \|\bar{G}_\lambda(\cdot,t)\|_p^2 \, dF_n(t) \le \int \|\bar{G}_\lambda(\cdot,t)\|_p^2 \, dF(t) + K_4 d_n \|\bar{G}_\lambda\|^2_{W_2^{p+d}(\Omega\times\Omega)}$$

Now Lemma 4.1, inequality (4.9a), and Assumption 3 imply the existence of constants $C_i = C_i(\Omega,m,p,F)$, $i = 1,2$, for which

$$\int \|\bar{G}_\lambda(\cdot,t)\|_p^2 \, dF(t) \le C_1 \lambda^{-(2p+d)/2m}$$

and

$$\|\bar{G}_\lambda\|^2_{W_2^{p+d}(\Omega\times\Omega)} \le C_2 \lambda^{-(2p+3d)/2m} .$$

Hence

$$\int \|\bar{G}_\lambda(\cdot,t)\|_p^2 \, dF_n(t) \le C_1 \lambda^{-(2p+d)/2m} + K_4 C_2 d_n \lambda^{-(2p+3d)/2m}$$

If we set $p = d$ in this latter result, and substitute the bounds back into (4.10), then the desired result is obtained.

#

Theorem 4.3. Suppose $m > 3d/2$, and that Assumptions 2,3, and 4 hold.

(i) There exists $n_0 = n_0(\Omega,m,\{\lambda_n\},F,\{d_n\})$ such that for all $n \geq n_0$ and all j, $1 \leq j \leq n$,

$$g_{n\lambda j} = \sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu} g_{\lambda j} ,$$

where the series converges in $W_2^p(\Omega)$ for any $p \in Z_+$, $p < 2m - 3d/2$. Moreover, for any fixed $n \geq n_0$, this convergence is uniform in j $(1 \leq j \leq n)$ and $\lambda \in [\lambda_n,\infty)$.

(ii) There exists $n_1 = n_1(\Omega,m,\{\lambda_n\})$ such that for all $n \geq n_1$ and any $p \in Z_+$, $p < 2m - 3d/2$,

$$\|g_{n\lambda k} - g_{\lambda k}\|_p \leq K_7(d_n\lambda^{-5d/4m})\ \lambda^{-(p-d)/2m}\|g_{\lambda k}\|_d$$

where $K_7 = K_7(\Omega,m,p,F)$ is a constant.

<u>Proof.</u>(i) For any fixed $t_0 \in \Omega$, let $\bar{G}_\lambda(t,\tau) = G_\lambda(t,\tau)/f(\tau)$ as before, and put

$$g_{\lambda 0} = \bar{G}_\lambda(\cdot,t_0) .$$

If $\nu \in Z_+$ and $0 \leq p < 2m - 3d/2$, then Lemma 4.2 (iii) yields

$$\|R_{n\lambda}^{\nu} g_{\lambda 0}\|_p \leq K_5(K_6 d_n\lambda^{-5d/4m})^{\nu}\ \lambda^{-(p-d)/2m}\|g_{\lambda 0}\|_d \tag{4.11}$$

Hence, if only

$$d_n\lambda_n^{-5d/4m} < K_6^{-1} \tag{4.12}$$

then the series in the statement of the Lemma converges in $W_2^p(\Omega)$. By Assumption 2, there is an $n_0 = n_0(\Omega,m,\{\lambda_n\},F,\{d_n\})$ for which (4.12) holds for all $n \geq n_0$. Lemma 4.1 (iii) and (4.9a) give the following bound, independent of t_0:

$$\|g_{\lambda 0}\|_d \leq K_3\lambda^{-d/m},$$

which implies the convergence of the series uniformly in j. Substituting this latter inequality back into (4.11) and putting $C_4 = K_5K_3$ yields

$$\|R_{n\lambda}^{\nu} g_{\lambda 0}\|_p \leq C_4(K_5 d_n\lambda^{-5d/4m})^{\nu}\ \lambda^{-(p+d)/2m} \tag{4.13}$$

which implies the convergence uniformly in $\lambda \geq \lambda_n$.

To complete the proof of (i), we need only show that

$$h \equiv \sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu} g_{\lambda k} \tag{4.14}$$

is in fact equal to $g_{n\lambda k}$. In order to do this, we will first show that for all $\chi \in W_2^m(\Omega)$, and any fixed to $t_o \in \Omega$,

$$\lambda(\bar{G}_\lambda(t_0,\cdot),\chi)_m + \int \bar{G}_\lambda(t_0,\cdot)\chi dF = \chi(t_0) \ . \tag{4.15}$$

We claim that $N_m^{2m}(\Omega)$ is dense in $W_2^m(\Omega)$, so that is suffices to prove (4.15) for $\chi \in N_m^{2m}$. This density claim follows since the linear span of the eigenfunctions $\{\phi_\nu\}$ is dense in both N_m^{2m} and N_m^m, and the latter is the same as W_2^m but with an equivalent norm (Proposition 3.1). Now for $\chi \in N_m^{2m}$, we may apply the Green's formula and note that the boundary terms vanish to obtain that the left hand side of (4.15) is equal to

$$\int_\Omega \{\lambda f(\tau)^{-1}[(-\Delta)^m \chi](\tau) + \chi(\tau)\}\, G_\lambda(t_o,\tau)d\tau \ .$$

However, this last quantity is equal to $\chi(t_0)$ by the definition of G_λ as a Green's function.

Now note that

$$h - g_{\lambda k} = \sum_{\nu=1}^{\infty} R_{n\lambda}^{\nu} g_{\lambda k} = R_{n\lambda} \sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu} g_{\lambda k} = R_{n\lambda} h \ ,$$

where the second equality is justified (if $m > 3d/2$) by the fact that the series converges absolutely in $W_2^m(\Omega)$, and $R_{n\lambda}$ is a continuous operator on $W_2^m(\Omega)$. Furthermore, using Fubini's theorem, we obtain for any $x \in W_2^m(\Omega)$ that

$$\lambda(R_{n\lambda}h,x)_m + \int (R_{n\lambda}h)x dF =$$

$$= \lambda \sum_{|\alpha|=m} \int \Big\{ \int G_\lambda^{\alpha,0}(t,\tau)h(\tau)d[F(\tau)-F_n(\tau)]\Big\}(D^\alpha x)(t)dt$$

$$+ \int \Big\{ \int \bar{G}_\lambda(t,\tau)h(\tau)d[F(\tau) - F_n(\tau)]\Big\}x(t)dF(t)$$

$$= \int \Big\{\lambda \sum_{|\alpha|=m} \int \bar{G}_\lambda^{\alpha,0}(t,\tau)(D^\alpha x)(t)dt + \int \bar{G}_\lambda(t,\tau)x(t)dF(t)\Big\}\cdot h(\tau)d[F(\tau)-F_n(\tau)]$$

$$= \int x(\tau)h(\tau)d[F(\tau) - F_n(\tau)] \ ,$$

where the last line follows from (4.15). These last two sets of displayed calculations can now be used to show that for any $x \in W_2^m(\Omega)$,

$$\begin{aligned}\lambda(h,x)_m &+ \int hx dF_n \\ &= \lambda(g_{\lambda k},x)_m + \int g_{\lambda k} x \, dF + \lambda(R_{n\lambda}h,x)_m + \int (R_{n\lambda}h)x \, dF - \int hx \, d[F-F_n] \\ &= x(t_k) ,\end{aligned}$$

where the last line follows from (4.15). It now follows from Proposition 2.1(i) that $h = g_{n\lambda k}$, and hence that the proof of (i) is complete.

(ii) If $n \geq n_0$, we have from (4.11) that

$$\begin{aligned}\|g_{n\lambda k} - g_{\lambda k}\|_p &\leq \sum_{\nu=1}^{\infty} \|R_{n\lambda}^{\nu} g_{\lambda k}\|_p \\ &\leq K_6 \lambda^{-(p-d)/2m} \|g_{\lambda k}\|_d \sum_{\nu=1}^{\infty} (K_5 d_n \lambda^{-5d/4m})^{\nu} \\ &= C_5 d_n \lambda^{-(p-d)/2m} \|g_{\lambda k}\|_d \, [1 - K_5 d_n \lambda^{-5d/4m}]^{-1} ,\end{aligned}$$

where $C_5 = C_5(\Omega,m,p,F)$. If we take $n_1 = n_1(\Omega,m,\{\lambda_n\},\{d_n\})$ sufficiently large that for all $n \geq n_1$,

$$d_n \lambda_n^{-5d/4m} \leq (2K_5)^{-1},$$

then part (ii) follows.

∎

5. Main Theorem. The results of the previous sections are brought together here to prove the following main result.

Theorem 5.1. Suppose Assumptions 1 through 4 hold, and that $m > (3d/2)$. Let $p \in Z_+$, with

$$p < 2m - (3d/2) .$$

Suppose $g \in N_m^q(\Omega)$ where q satisfies

$$d < q \le 2m + p, \quad p < q .$$

If $g_{n\lambda} = S_{n\lambda} \underset{\sim}{z}$ with the observational model (1.1), then as $n \to \infty$

$$E\|g_{n\lambda} - g\|_p^2 = O\left[\lambda^{(q-p)/m} + n^{-1}\lambda^{-(2p+d)/2m}\right]$$

uniformly in $\lambda \in [\lambda_n, \Lambda_n]$.

Remark 5.2. The last phrase means there exists an n_0 (depending only on $\Omega, m, p, q, \{\lambda_n\}$, $\{\Lambda_n\}$, and $\{d_n\}$) such that there is a constant C for which $n \ge n_0$ implies for all $\lambda \in [\lambda_n, \Lambda_n]$,

$$E\|g_{n\lambda} - g\|_p^2 \le C\left[\lambda^{(q-p)/m} + n^{-1}\lambda^{-(2p+d)/2m}\right] .$$

Remark 5.3. From the two lemmas below, we see that a somewhat sharper result can be stated if one is willing to use the equivalent $N_m^p(\Omega)$ norm, namely, for all $\theta > 0$,

$$\sup_{\|g\|_{N_m^q} = \theta} E\|g_{n\lambda} - g\|^2_{N_m^p(\Omega)} \sim A_{pq}(\lambda)\theta^2 + \int \|G_\lambda(\cdot,t)\|^2_{N_m^p(\Omega)}\, dF(t) .$$

Here, $A_{pq} : R_+ \to R_+$ is given in (5.2), and satisfies the relation $A_{pq}(\lambda) = \lambda^{(q-p)/m}$. The supremum is taken over all $g \in N_m^q(\Omega)$ for which $\|g\|_{N_m^q(\Omega)} = \theta$.

Proof of Theorem. Using Assumption 1, we calculate

$$\begin{aligned} E\|g_{n\lambda} - g\|_p^2 &= E\|S_{n\lambda}\underset{\sim}{g}(\Delta_n) + \underset{\sim}{\varepsilon} - g\|_p^2 \\ &= \|S_{n\lambda}\underset{\sim}{g}(\Delta_n) - g\|_p^2 + 2n^{-1}\sum_k (E\varepsilon_k)\, \langle S_{n\lambda}\underset{\sim}{g}(\Delta_n) - g, g_{n\lambda k}\rangle_p + E\|S_{n\lambda}\underset{\sim}{\varepsilon}\|_p^2 \\ &= \|S_{n\lambda}\underset{\sim}{g}(\Delta_n) - g\|_p^2 + E\|S_{n\lambda}\underset{\sim}{\varepsilon}\|_p^2 . \end{aligned}$$

This is the familiar decomposition of mean squared error into bias squared plus variance. The theorem is now an immediate consequence of the following two lemmas and the equivalence of $\|\cdot\|_p$ and $\|\cdot\|_{N_m^p(\Omega)}$ (Proposition 3.1).

__Lemma 5.4.__ Under Assumptions 2,3, and 4, if $m > (3d/2)$, $p \in Z_+$ and $q \in [d,\infty)$ satisfy $p < 2m - (3d/2)$ and $p < q \leq 2m + p$, then for every $\theta > 0$,

$$\sup \{\|S_{n\lambda}g(\Delta_n) - g\|^2_{N_m^p(\Omega)} : g \in N_m^q(\Omega) \text{ and } \|g\|_{N_m^q(\Omega)} = \theta\} \sim A_{pq}(\lambda)\theta^2$$

as $n \to \infty$. Here, the function A_{pq} is determined by $p,q,\Omega,m,$ and F, and furthermore

$$A_{pq}(\lambda) \asymp \lambda^{(q-p)/m} .$$

For both asymptotic relations, there exists $n_0 = n_0(\Omega,m,p,q,\{\lambda_n,\Lambda_n\},\{d_n\}, F)$ such that for all $n \geq n_0$, the relations hold uniformly in $\lambda \in [\lambda_n,\Lambda_n]$.

__Proof.__ First note that

$$S_{n\lambda}g(\Delta_n) = n^{-1}\sum_{k=1}^{n} g(t_k)g_{n\lambda k} = \sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu} \int \bar{G}_\lambda(\tau,\cdot)g(\tau)dF_n(\tau) ,$$

so that

$$S_{n\lambda}g(\Delta_n) - g = [\sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu} (\int \bar{G}_\lambda(\tau,\cdot)g(\tau)dF(\tau) - R_{n\lambda}g)] - g = \sum_{\nu=0}^{\infty} R_{n\lambda}^{\nu}\varepsilon_\lambda ,$$

where

$$\varepsilon_\lambda = \int G_\lambda(\cdot,\tau)g(\tau)d\tau - g .$$

In order to show convergence of the series and obtain bounds on the terms, we need to estimate ε_λ. Now

$$\varepsilon_\lambda = \sum_{\nu=1}^{\infty} (1 + \lambda\gamma_\nu)^{-1}\lambda\gamma_\nu(g,\phi_\nu)_0\phi_\nu ,$$

where $\{\gamma_\nu\}$, $\{\phi_\nu\}$ is the eigensystem for $\mathcal{D}$ introduced above. We see immediately that

$$\|\varepsilon_\lambda\|^2_{N_m^p(\Omega)} = \sum_{\nu=1}^{\infty} (1 + \gamma_\nu)^{p/m}(1 + \lambda\gamma_\nu)^{-2}(\lambda\gamma_\nu)^2(g,\phi_\nu)_0^2 \leq A_{pq}(\lambda)\|g\|^2_{N_m^q(\Omega)} \tag{5.1}$$

where

$$(5.2) \qquad A_{pq}(\lambda) = \sup_{\nu>0} (1+\gamma_\nu)^{(p-q)/m}(1+\lambda\gamma_\nu)^{-2}(\lambda\gamma_\nu)^2 .$$

Moreover, if the sup is attained at some (finite) ν_0, then the upper bound in (5.1) is attained at any scalar multiple of ϕ_{ν_0}. Utilizing the eigenvalue estimate (4.3) gives

$$(5.3) \qquad A_{pq}(\lambda) \le C_2 \sup_{\nu>0} \nu^{2(p-q)/d}(1+C_1\lambda\nu^{2m/d})(\lambda\nu^{2m/d})^2$$

$$\le C_2\lambda^{(q-p)/m} \sup_{x>0} (1+C_1 x)^{-2}x^{(2m+p-q)/m} ,$$

where $C_i = C_i(\Omega,m,p,q,F)$, $i = 1,2$, are constants. Here we have replaced the discrete variable $\lambda\nu^{2m/d}$ by the continuous variable x. Now if

$$p < q \le 2m + p$$

then as $x \to \infty$,

$$(1+C_1x)^{-2}x^{(2m+p-q)/m} \to 0 ,$$

and remains bounded as $x \to 0$. This shows that the sup in (5.2) is attained at some ν_0, and hence that the upper bound in (5.1) is attained. Note that $(1+C_1x)^{-2}x^{(2m+p-q)/m}$ remains finite if only $q \ge p$. Furthermore, one can obtain a lower bound for $A_{pq}(\lambda)$ which is of the same form as (5.3) (only the constants are different), and so

$$A_{pq}(\lambda) \approx \lambda^{(q-p)/m} .$$

To complete the proof of the Lemma, we need only show that $\sum_{\nu=1}^{\infty} R^\nu_{n\lambda}\epsilon_\lambda$ is asymptotically negligible compared to ϵ_λ. We apply Lemma 4.2(iii) to obtain for all $\nu \ge 1$

$$\|R^\nu_{n\lambda}\epsilon_\lambda\|_p \le C_3(C_2 d_n\lambda^{-5d/4m})^\nu\lambda^{-(p-d)/2m}\|\epsilon_\lambda\|_d ,$$

and hence that for all n sufficiently large

$$\|\sum_{\nu=1}^{\infty} R^\nu_{n\lambda}\epsilon_\lambda\|_p \le K(d_n\lambda^{-5d/4m})\lambda^{(d-p)/2m}\|\epsilon_\lambda\|_d$$

$$\le K'\|g\|_q(d_n\lambda^{-5d/4m})\ \lambda^{(d-p)/2m}\ \lambda^{(q-d)/2m}$$

$$\le K'\|g\|_q(d_n\lambda^{-5d/4m})\lambda^{(q-p)/2m}$$

where K, K' are constants depending on $\Omega, m, p, q,$ and F. The fact that $q > d$ was used at the second step. By the first part of the proof, Assumption 2 implies the last expression is asymptotically negligible compared with $\|\varepsilon_\lambda\|_p$, completing the proof. #

__Lemma 5.5.__ Let $m > 3d/2$, and suppose Assumptions 1,2,3, and 4 hold. If $p \in Z_+$ satisfies

$$0 \leq p < 2m - (3d/2) ,$$

then

$$E\|S_{n\lambda}\underset{\sim}{\varepsilon}\|_p^2 \approx n^{-1}\lambda^{-(2p+d)/2m}$$

as $n \to \infty$. Furthermore, there exists an $n_0 = n_0(\Omega, m, F, p, \{\lambda_n, \Lambda_n\}, \{d_n\})$ such that the constants may be chosen independently of $\lambda \in [\lambda_n, \Lambda_n]$ for all $n \geq n_0$.

__Proof.__ Since the errors are mean zero and uncorrelated, we obtain

$$(5.4) \qquad E\|S_{n\lambda}\underset{\sim}{\varepsilon}\|_p^2 = \sigma^2 n^{-2} \sum_{k=1}^{n} \|g_{n\lambda k}\|_p^2$$

$$= \sigma^2 n^{-2} \sum_{k=1}^{n} \{\|g_{\lambda k}\|_p^2 + 2 < g_{\lambda k}, g_{\lambda k} - g_{n\lambda k} >_p + \|g_{\lambda k} - g_{n\lambda k}\|_p^2\} .$$

We now deal with each of the three terms in turn.

Firstly, note that

$$n^{-1} \sum_k \|g_{\lambda k}\|_p^2 = \int \|\bar{G}_\lambda(t,\cdot)\|_p^2 dF_n(t)$$

$$= \int \|\bar{G}_\lambda(t,\cdot)\|_p^2 dF(t) - \int \|\bar{G}_\lambda(t,\cdot)\|_p^2 d[F(t) - F_n(t)] .$$

Using Assumption 3 and Lemma 4.1(i), we obtain

$$(5.5) \qquad \int \|\bar{G}_\lambda(t,\cdot)\|_p^2 dF(t) \approx \lambda^{-(2p+d)/2m} .$$

Also, Lemma 4.2(i) and the assumption that $p < 2m - (3d/2)$ yield

$$(5.6) \qquad \int \|\bar{G}_\lambda(t,\cdot)\|_p^2 d[F(t) - F_n(t)] \leq K_4 d_n \|\bar{G}_\lambda\|^2_{W_2^{p+d}(\Omega\times\Omega)} = d_n O\{\lambda^{-(2p+3d)/2m}\}$$

$$= \{d_n \lambda^{-5d/4m}\}\lambda^{d/4m} O\{\lambda^{-(2p+d)/2m}\} = o\{\lambda^{-(2p+d)/2m}\} .$$

Assumption 2 is used at the last step.

Turning to the second term in (5.4), an application of Cauchy-Schwarz gives

$$| n^{-1} \sum_k \langle g_{\lambda k}, g_{\lambda k} - g_{n\lambda k} \rangle_p | \le n^{-1} \sum_k \|g_{\lambda k}\|_p \|g_{\lambda k} - g_{n\lambda k}\|_p$$

$$\le K_7 (d_n \lambda^{-5d/4m}) \lambda^{-(p-d)/2m} n^{-1} \sum_k \|g_{\lambda k}\|_p \|g_{\lambda k}\|_d$$

where Theorem 4.3(ii) was used at the last step. Now by Lemma 4.2(iv),

$$\lambda^{-(p-d)/2m} n^{-1} \sum_k \|g_{\lambda k}\|_p \|g_{\lambda k}\|_d = [1 + d_n \lambda^{-d/m}] O(\lambda^{-(2p+d)/2m}) \; .$$

Combining the last two displayed estimates and using Assumption 2 gives

$$(5.7) \qquad n^{-1} \sum_k \langle g_{\lambda k}, g_{\lambda k} - g_{n\lambda k} \rangle_p = o(\lambda^{-(2p+d)/2m}) \; .$$

Finally, the third term in (5.4) is easily dealt with in the same manner, and one obtains

$$(5.8) \qquad n^{-1} \sum_k \|g_{\lambda k} - g_{n\lambda k}\|_p^2 = (d_n \lambda^{-5d/4m})^2 \, (\lambda^{-(2p+d)/2m}) = o(\lambda^{-(2p+d)/2m}) \; .$$

Combining (5.5), (5.6), (5.7), and (5.8), and inserting them into (5.4) yields

$$E\|S_{n\lambda}\varepsilon\|_p^2 \sim \sigma^2 n^{-1} \int \|G_\lambda(\cdot,t)\|_p^2 dF(t) \approx n^{-1} \lambda^{-(2p+d)/2m} \; .$$

The statement regarding uniformity follow from the corresponding statements in Lemmas 4.1, 4.2, and Theorem 4.3.

#

Remarks 5.6. throughout these remarks, we adopt the notation and hypotheses of Theorem 5.1.

(i) The best possible rate of convergence implied by Theorem 5.1 is always obtained if

$$\lambda = n^{-2m/(2q+d)}$$

in which case

$$E\|g_{n\lambda} - g\|_p^2 = O\{n^{-2(q-p)/(2q+d)}\} \; .$$

As will be noted in (vi) below, implicit here are certain lower bounds on q besides those stated in the theorem. We proceed to analyze some of the various cases with regard to smoothness of g and boundary conditions.

(ii) If $q < m + 1/2$, then $N_m^q(\Omega) = B_{22}^q(\Omega)$, and there is no difficulty with boundary conditions.

(iii) If $m + 1/2 \leq q < 3m + 1/2$, then necessarily $B_k g = 0$ on $\partial\Omega$ for all $k \in Z_+$, $1 \leq k \leq m$, such that $m + k - 1 \leq q - 1/2$ (see Proposition 3.1). For a particular p, the requirement $q \leq 2m + p$ may be the limiting factor. In particular, if $p = 0$, then there is no gain in convergence rate for q beyond $2m$. This is an example of the saturation phenomenon familiar in approximation theory. Note however that we can still make gains in estimating higher order derivatives $(p \geq 1)$.

(iv) If $3m + 1/2 \leq q \leq 2m + p < 4m - (3d/2)$, then not only does g satisfy the natural boundary conditions $B_k g = 0$ on $\partial\Omega$, $1 \leq k \leq m$, but also some second order natural boundary conditions, namely $B_k \mathcal{D} g = 0$ on $\partial\Omega$ for k such that $3m + k - 1 \leq q - 1/2$. This, of course, only has an effect on estimating derivatives of order $m + 1, m + 2, \ldots, [2m - (3d/2)]$.

(v) Note that when $q \geq m + 1/2$, the assumption $g \in N_m^q(\Omega)$ limits g in both smoothness and boundary conditions. If, for example, $g \in W_2^{2m}(\Omega)$, but for some j, $1 \leq j < m$, we have $B_k g = 0$ on $\partial\Omega$ if $k < j$ but $B_j g \not\equiv 0$ on $\partial\Omega$, then $g \in N_m^q(\Omega)$ for all $q < m + j - 1/2$, but no larger q. Hence, we have for every $\varepsilon < 0$ that the rate

$$E\|g_{n\lambda} - g\|_p^2 = O\{n^{-2(m+j-1/2-\varepsilon)/(2m+2j-1-2\varepsilon+d)}\}$$

is obtainable. It would be interesting to sharpen this result, by going to logarithms for example.

(vi) In order to verify Assumption 2, it is necessary to know how fast $d_n = \sup|F - F_n|$ can go to zero. For $d = 1$ and $\Omega = (0,1)$, one can easily check that

$$d_n \geq (2n)^{-1}$$

with equality just in case

$$t_k = (k - 1/2)/n \qquad 1 \leq k \leq n .$$

For general d, if $n = j^d$ for some $j \in Z_+$, then the uniform discrete measure concentrated on the grid of the form

$$j^{-1}(k_1 - \tfrac{1}{2}, \ldots, k_d - \tfrac{1}{2}) \quad , \quad 1 \leq k_i \leq j$$

approaches uniform (Lebesque) measure on the cube $(0,1)^d$ in Kolmogorov metric at the rate $j^{-1} = n^{-1/d}$. Under Assumption 3, F looks like uniform measure on sufficiently small cubes, so the rate

$$d_n = n^{-1/d}$$

is at least attainable. However, if one chooses the knots randomly (i.e. if $t_1, t_2, \ldots$ are independent and identically distributed with the distribution F), then the rate can be improved to

$$d_n = O_p(n^{-1/2}) ,$$

where O_p means "big oh in probability," See equation (2.4.3) in Gaenssler and Stute [6]. We conjecture that $d_n \approx n^{-1}$ is obtainable in any dimension, but the construction of knot sequences attaining this rate appears to be a nontrivial problem. In general, if

$$d_n = n^{-r}$$

for some $r > 0$, and if also

$$\lambda = n^{-2m/(2q+d)} ,$$

(so that the optimal rate is obtained), then Assumption 2 requires that

$$q > d(5-2r)/(4r) .$$

This is a stronger assumption than the minimal one required for Theorem 5.1 (i.e. that $q > d$) unless $r > 5/6$.

#

6. Extension to Larger Domains. Suppose that F and Ω satisfy Assumptions 3 and 4, and that Ω' is a domain satisfying

$$\Omega \subseteq \Omega' ,$$

e.g. $\Omega' = \mathbb{R}^d$. Consider the smoothing spline estimate $g'_{n\lambda}$ obtained by minimizing over $h \in W_2^m(\Omega')$ the quadratic form

$$\frac{1}{n} \sum_k (z_k - h(t_k))^2 + \lambda \sum_{|\alpha|=m} \int_{\Omega'} (D^\alpha h(t))^2 dt .$$

The only difference between $g'_{n\lambda}$ and $g_{n\lambda}$ is that the objective function involves $L_2(\Omega')$ norms of m^{th} order derivatives rather than $L_2(\Omega)$ norms. If $\Omega' = \mathbb{R}^d$, then $g'_{n\lambda}$ is called a "thin plate" smoothing spline [22], or Laplacian smoothing spline [24]. Our convergence rates will be in terms of the expected value of the square of the $L_2(F_n)$ seminorm of the error, viz.

$$T'_{n\lambda} = E \left[\int (g'_{n\lambda}-g)^2 dF_n \right]$$

$$= E \left[n^{-1} \sum_{k=1}^{n} (g'_{n\lambda}(t_k) - g(t_k))^2 \right] ,$$

with $T_{n\lambda}$ defined by replacing $g'_{n\lambda}$ with $g_{n\lambda}$.

We define an $n \times n$ matrix $A'_{n\lambda}$ by

(6.1) $$A'_{n\lambda} \underset{\sim}{z} = (S'_{n\lambda}\underset{\sim}{z})(\Delta_n),$$

where $S'_{n\lambda} : \mathbb{R}^n \to W_2^m(\Omega')$ is the smoothing spline operator for the domain Ω' (see Definition 1.1). It is easily checked that

$$T'_{n\lambda} = \int [S'_{n\lambda}(g(\Delta_n)) - g]^2 dF_n + n^{-1}\sigma^2 \mathrm{Tr}[(A'_{n\lambda})^2] .$$

See, for example, equation (1.7) of Craven and Wahba [5]. A simple argument (Lemma 4.1 of [5]) shows that if $g \in W_2^m(\Omega')$, then

$$\int [S'_{n\lambda}(g(\Delta_n)) - g]^2 dF_n \le \lambda \int_{\Omega'} (g^{(m)}(u))^2 du$$

$$= \lambda |g|^2_{W_2^m(\Omega')} ,$$

for all $n \geq 1$ and all $\lambda > 0$. Hence, to obtain an upper bound for $T'_{n\lambda}$,

it is only necessary to bound $\mathrm{Tr}(A'_{n\lambda})^2$. We shall do this by showing that the eigenvalues of $A'_{n\lambda}$ are bounded by the corresponding eigenvalues of $A_{n\lambda}$, the matrix obtained by deleting the primes in (6.1). A simple variant of the proof of Lemma 5.5 then gives an upper bound on $\mathrm{Tr}\, A^2_{n\lambda}$, assuming that F and Ω satisfy the assumptions.

Define the quadratic form

$$B_n(u,v) = \int uv\,dF_n$$

for either $u,v \in W_2^m(\Omega)$ or $u,v \in W_2^m(\Omega')$.

We assume that

$$\Delta_n \subseteq \Omega \subseteq \Omega' .$$

Also define the quadratic forms

$$A_{n\lambda}(u,v) = B_n(u,v) + \lambda(u,v)_{W_2^m(\Omega)} ,$$

$$A'_{n\lambda}(u,v) = B_n(u,v) + \lambda(u,v)_{W_2^m(\Omega')} ,$$

with domains $W_2^m(\Omega)$ and $W_2^m(\Omega')$, respectively. It follows from the proof of Proposition 2.1 that $A_{n\lambda}$ and $A'_{n\lambda}$ are strictly positive definite. Furthermore, the codimension of the null space of B_n is n for both domains, and the Rayleigh quotients $B_n(u,u)/A_{n\lambda}(u,u)$ and $B_n(u,u)/A'_{n\lambda}(u,u)$ are bounded by 1. Hence, there are n positive eigenvalues.

(6.2) $$\mu_{n\lambda 1} \geq \mu_{n\lambda 2} \geq \dots \geq \mu_{n\lambda n} > 0$$

for the Rayleigh quotient $B_n(u,u)/A_{n\lambda}(u,u)$. Here, the eigenvalues are defined recursively by

$$\mu_{n\lambda 1} = \sup \{B_n(u,u) : u \in W_2^m(\Omega),\ A_{n\lambda}(u,u) = 1\} ,$$

where the supremum is obtained at $u_{n\lambda 1}$, and assuming $(\mu_{n\lambda j}, u_{n\lambda j})$ has been defined for $1 \leq j \leq k-1$,

$$\mu_{n\lambda k} = \sup \{B_n(u,u) : u \in W_2^m(\Omega),\ A_{n\lambda}(u,u) = 1, \text{ and } A_{n\lambda}(u,u_{n\lambda j}) = 0 \text{ for } 1 \leq j \leq k-1\} .$$

See Theorem 3.2.3 of Weinberger [23]. Similary, we let

(6.3) $$\mu'_{n\lambda 1} \geq \mu'_{n\lambda 2} \geq \dots \geq \mu'_{n\lambda n} > 0$$

be the eigenvalues for the Rayleigh quotient $B_n(u,u)/A'_{n\lambda}(u,u)$. For a more complete account of these eigensystems and further applications (in one dimension) see Speckman [14], Section 5. The most important fact for our purposes is that (6.2) (respectively (6.3)) are the eigenvalues for the matrix $A_{n\lambda}$ (respectively $A'_{n\lambda}$). To see this, note that the variational equation of Proposition 2.1 may be written as

$$A_{n\lambda}(v,g_{n\lambda}) = B_n(v,\zeta) \quad \forall\, v \in W_2^m(\Omega)$$

where $\zeta \in W_2^m(\Omega)$ is any function satisfying

$$\underset{\sim}{\zeta}(\Delta_n) = \underset{\sim}{z}\ .$$

Hence, if for some $\gamma > 0$ we have

$$A_{n\lambda}\,\underset{\sim}{z} = \gamma \underset{\sim}{z}$$

then we may take $\zeta = \gamma^{-1} g_{n\lambda}$ so that

$$A_{n\lambda}(v,g_{n\lambda}) = \gamma^{-1} B_{n\lambda}(v,g_{n\lambda}),\ \forall\, v \in W_2^m(\Omega)\ ,$$

from which it follows that γ is an eigenvalue of the Rayleigh quotient $B_n/A_{n\lambda}$.

Lemma 6.1. Let $m > d/2$ and suppose that $\Omega \subseteq \Omega'$. Then the eigenvalues in (6.2) and (6.3) satisfy

$$\mu'_{n\lambda k} \le \mu_{n\lambda k} \quad \text{for} \quad 1 \le k \le n\ .$$

Proof. The assumption $m > d/2$ implies that $W_2^m(\Omega)$ and $W_2^m(\Omega')$ may be embedded in $C(\Omega)$ and $C(\Omega')$, respectively (Theorem 4.6.1(e) of Triebel [17]). Hence, B_n, $A_{n\lambda}$, and $A'_{n\lambda}$ are well defined, and so is the restriction operator $R : W_2^m(\Omega') \to W_2^m(\Omega)$, i.e. $(Ru)(t) = u(t)$ for all $t \in \Omega$, all $u \in W_2^m(\Omega')$.

Then the following inequalities are obvious:

$$B_n(Ru,Ru) \ge \left(\frac{B_n(u,u)}{A'_{n\lambda}(u,u)}\right) A_{n\lambda}(Ru,Ru)$$

$$A_{n\lambda}(Ru,Ru) \ge \frac{B_n(u,u)}{A'_{n\lambda}(u,u)} A'_{n\lambda}(u,u)$$

for all $u \in W_2^m(\Omega')/\{0\}$. The lemma now follows from the "Mapping Principle," Theorem 3.6.1 of Weinberger [23]. ∎

Theorem 6.2. Suppose Assumptions 1,2,3, and 4 hold (for the domain Ω), and that $m > 3d/2$. Then, for any domain $\Omega' \supseteq \Omega$, there exists a finite constant $C > 0$ depending on m, Ω, Ω', and F, and an integer n_0 such that for all $n \geq n_0$ and all $\lambda \in [\lambda_n, \Lambda_n]$,

$$T'_{n\lambda} \leq \lambda |g|^2_{W_2^m(\Omega')} + c\, \sigma^2 n^{-1} \lambda^{-d/2m} .$$

Proof. In view of Lemma 6.1 and the remarks preceding it, it suffices to show that

$$\mathrm{Tr}\, A^2_{n\lambda} = O(\lambda^{-d/2m}) ,$$

uniformly in $\lambda \in [\lambda_n, \Lambda_n]$ as $n \to \infty$. Since

$$\begin{aligned} n^{-1} \sigma^2\, \mathrm{Tr}\, A^2_{n\lambda} &= E\left[n^{-1} \sum_{j=1}^{n} (S_{n\lambda}\varepsilon)^2(t_j)\right] \\ &= E\|S_{n\lambda}\varepsilon\|^2_{L_2(F_n)} , \end{aligned}$$

and by Assumption 4 and Lemma 5.5,

$$\begin{aligned} E\|S_{n\lambda}\varepsilon\|^2_{L_2(F)} &= E\|S_{n\lambda}\varepsilon\|^2_0 \\ &= n^{-1}\lambda^{-d/2m} , \end{aligned}$$

uniformly in $\lambda \in [\lambda_n, \Lambda_n]$, it suffices to prove that

$$E\left[\|S_{n\lambda}\varepsilon\|^2_{L_2(F_n)} - \|S_{n\lambda}\varepsilon\|^2_{L_2(F)}\right] = o(n^{-1}\lambda^{-d/2m}) , \tag{6.4}$$

uniformly in $\lambda \in [\lambda_n, \Lambda_n]$. To this end, apply Lemma 4.2 (i) and Lemma 5.5 to obtain

$$\begin{aligned} \left| E \int (S_{n\lambda}\varepsilon)^2 d[F - F_n] \right| &\leq K_4 d_n\, E\|S_{n\lambda}\varepsilon\|^2_d \\ &= d_n n^{-1} \lambda^{-3d/2m} \\ &= (d_n \lambda^{-d/m})(n^{-1}\lambda^{-d/2m}) . \end{aligned}$$

Equation (6.4) now follws from Assumption 2.

REFERENCES

[1] Adams, R.A. (1975) Sobolev Spaces, Academic Press, New York.

[2] Agmon, S. (1965) Lectures on elliptic boundary value problems, Van Nostrand, New York.

[3] Cogburn, R. & Davis, H. (1974), Periodic splines and spectral estimation, Ann. Statist. 2,1108 - 1126.

[4] Cox, D. (1981), Asymptotics for M-type smoothing splines, Tech. report no. 654, Dept. of Statistics, Univ. of Wisconsin, Madison.

[5] Craven, P . & Wahba, G. (1979), Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross validation, Numerische Math. 31, 377-403.

[6] Gaenssler, P. & Stute, W. (1979), Empirical processes: a survey of results for independently and identically distributed random variables, Ann. Probability 7, 193-243.

[7] Kimeldorf, G. & Wahba G. (1970), A correspondence between Bayesian estimation on stochastic processes and smoothing by splines, Ann. Math. Statist. 41, 495-502.

[8] Lions, J. & Magenes, E. (1972) Non-homogeneous boundary value problems and applications, vol. I, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 181, Springer-Verlag, Berlin.

[9] Ragozin, D. (1981), Error bounds for derivative estimates based on spline smoothing of exact or noisy data, preprint, Dept of Math., Univ. of Washington, Seattle.

[10] Ragozin, D., Powers, R., & Tong, T. (1981), Numerical differentiation using regularization and generalized crossvalidation, preprint, Dept. of Math., Univ. of Washington, Seattle.

[11] Rice, J. & Rosenblatt, M. (1980), Boundary effects on the behavior of smoothing splines, preprint, Dept. of Math., Univ. of Calif. San Diego.

[12] Silverman, B. (1981), On the estimation of a probability density function by the maximum penalized likelihood method, Math. Research Center Tech. report no. 2228, Univ. of Wisconsin, Madison.

[13] Speckman, P. (1982), Minimax estimates of linear functionals in a Hilbert space, to appear in Ann. Statist. 10.

[14] Speckman, P. (1981a), The asymptotic integrated mean square error for smoothing noisy data by splines, preprint, Dept. of Math., Univ. of Oregon, Eugene.

[15] Speckman, P. (1981b), Spline smoothing and optimal rates of convergence in nonparametric regression models, preprint, Dept. of Math., Univ. of Oregon, Eugene.

[16] Stone, C. (1980), Optimal rates of convergence for nonparametric estimators, Ann. Statist. 8, 1348-60.

[17] Triebel, H. (1978), Interpolation theory, function spaces, differential operators, North-Holland, New York.

[18] Utreras, F. (1980a), Sur le choix du parametre d'adjustment dans le lissage par functions spline, Numerische Math. 34, 15-28

[19] Utreras, F. (1980b), Natural spline functions, the associated eigenvalue problem, preprint from Universidad de Chile.

[20] Wahba, G. (1975), Smoothing noisy data with spline functions, Numerische Math. 24, 383-393.

[21] Wahba, G. (1979), How to smooth curves and surfaces with splines and crossvalidation, Tech. report no. 555, Dept. of Statistics, Univ. of Wisconsin, Madison.

[22] Wahba, G. (1979), Convergence rates of "thin plate" smoothing splines when the data are noisy, in Smoothing techniques for curve estimation, edited by T. Gasser and M. Rosenblatt, Lecture notes in math. no. 757, Springer-Verlag, New York.

[23] Weinberger, H. (1974), Variational methods for eigenvalue approximation, S.I.A.M. Regional Conf. Series in App. Math. no. 15, J. W. Arrowsmith Ltd., Bristol, England.

[24] Wendelberger, J. (1981), The computation of Laplacian smoothing splines with examples, Tech report no. 648, Dept. of Statistics, Univ. of Wisconsin, Madison.

DDC/jgb

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER #2437	2. GOVT ACCESSION NO. AD-A124	3. RECIPIENT'S CATALOG NUMBER 368
4. TITLE (and Subtitle) Convergence Rates for Multivariate Smoothing Spline Functions		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Dennis D. Cox		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS (see Item 18 below)		12. REPORT DATE October 1982
		13. NUMBER OF PAGES 39
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

* Sponsored by:

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Smoothing spline, nonparametric regression, rates of convergence, Sobolev norms, elliptic boundary value problems.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Given data $z_i = g(t_i) + \varepsilon_i$, $1 \le i \le n$, where g is the unknown function, the t_i are known d-dimensional variables in a domain Ω, and the ε_i are i.i.d. random errors, the smoothing spline estimate $g_{n\lambda}$ is defined to be the minimizer over h of $n^{-1}\Sigma(z_i-h(t_i))^2 + \lambda J_m(h)$, where $\lambda > 0$ is a smoothing parameter and $J_m(h)$ is the sum of the integrals over Ω of the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (Continued)

squares of all the m^{th} order derivatives of h. Under the assumptions that Ω is bounded and has a smooth boundary, $\lambda \to 0$ appropriately, and the t_i become dense in Ω as $n \to \infty$, bounds on the rate of convergerence of the expected square of $p^{\underline{th}}$ order Sobolev norm (L_2 norm of p^{th} derivatives), are obtained. These extend known results in the one dimensional case. The method of proof utilizes an approximation to the smoothing spline based on a Green's function for a linear elliptic boundary value problem. Using eigenvalue approximation techniques, these rate of convergence results are extended to faily arbitrary domains including $\Omega = R^d$, but only for the case $p = 0$, i.e. L_2 norm.